



## SEQUENCE LISTING

A <110> Black Jr., Charles A.

<120> COMPOSITIONS AND METHODS FOR ACTIVATING  
GENES OF INTEREST

<130> 5722-2(35722/191928)

<140> 09/446,402

<141> 1999-12-20

<150> PCT/US98/13093

<151> 1998-06-24

<150> 60/050,772

<151> 1997-06-25

<160> 19

<170> FastSEQ for Windows Version 4.0

<210> 1

<211> 4279

<212> DNA

<213> Artificial Sequence

<220>

<223> Recombinant Molecule containing multiple cloning  
site, kozak sequence, LacZ gene.

<221> misc\_feature

<222> (1)...(64)

<223> Multiple cloning site

<221> misc\_feature

<222> (65)...(79)

<223> Consensus sequence for the "Kozak sequence"  
(translation initiation)

<221> prim\_transcript

<222> (80)...(4279)

<223> Beta galactosidase

<400> 1

ttaatacgac	tcactatagg	ctagcctcga	gaattcacgc	gtggtacctc	tagagtcgac	60
ccggggccgcc	gccaccatgg	cgcagcacca	tggcctgaaa	taacctctga	aagagggaact	120
tggttaggta	ccttctgagg	cggaaagaac	cagctgtgga	atgtgtgtca	gttaggggtgt	180
ggaaaagtccc	caggctcccc	agcaggcaga	agtatgcaaa	gcatgcatct	caattagtca	240
gcaaccaggt	gtggaaaagtc	cccaggctcc	ccagcaggca	gaagtatgca	aagcatgcat	300
ctcaattagt	cagcaaccat	agtcccggcc	ctaactccgc	ccatcccggc	cctaactccg	360
cccagttccg	cccattctcc	gccccatggc	tgactaattt	tttttattta	tgcagaggcc	420
gaggccgcct	cggcctctga	gctattccag	aagtagtgag	gaggcttttt	tggaggccta	480
ggcttttgca	aaaagcttgg	gatctctata	atctcgcgca	acctattttc	ccctcgaaca	540

ctttttaagc	cgtagataaa	caggctggga	cacttcacat	gagcgaaaaa	tacatcgta	600
cctgggacat	gttgagatc	catgcacgta	aactcgcaag	ccgactgatg	ccttctgaac	660
aatggaaagg	cattattgcc	gtaagccgtg	gcggtctggt	accggtgggt	gaagaccaga	720
aacagcacct	cgaactgagc	cgcgatattg	cccagcggtt	caacgcgctg	tatggcgaga	780
tcgatcccg	cgttttacaa	cgtcgtgact	gggaaaaacc	tggcgttacc	caacttaatc	840
gccttgacgc	acatccccct	ttcgccagct	ggcgtaatag	cgaagaggcc	cgcaccgatc	900
gcccttccca	acagttgcgc	agcctgaatg	gcgaatggcg	ctttgcctgg	tttccggcac	960
cagaagcggg	gccggaaaagc	tggctggagt	gcgatcttcc	tgaggccgat	actgtcgtcg	1020
ttccctcaaa	ctggcagatg	cacggttacg	atgcgcccac	ctacaccaac	gtaacctatc	1080
ccattacggg	caatccgccc	tttgttccca	cggagaatcc	gacgggttgt	tactcgctca	1140
catttaaatg	tgatgaaagc	tggctacagg	aaggccagac	gcgaattatt	tttgatggcg	1200
ttactcggc	gtttcatctg	tgggtcaacg	ggcgctgggt	cggttacggc	caggacagtc	1260
gtttgccgct	tgaatttgac	ctgagcgcac	ttttacgcgc	cggagaaaaa	cgcctcgcgg	1320
tgatgggtgct	gcgttggaag	gacggcaggt	atctggaaga	tcaggatatg	tggcggaatga	1380
gcggcatttt	ccgtgacgct	tcgttgctgc	ataaaccgac	tacacaaaatc	agcgattttcc	1440
atgttgccac	tcgctttaat	gatgatttca	gccgcgctgt	actggaggct	gaagttcaga	1500
tgtgcggcga	gttgcgtagc	tacctacggg	taacagtttc	tttatggcag	ggtgaaacgc	1560
aggctgccag	cggcaccgcg	cctttcggcg	gtgaaattat	cgatgagcgt	ggtggttatg	1620
ccgatcgcgt	cacactacgt	ctgaacgctg	aaaaccgcaa	actgtggagc	gccgaaatcc	1680
cgaatctcta	tcgtgcggtg	gttgaaactgc	acaccgccga	cggcacgctg	attgaagcag	1740
aagcctgcga	tgctcggtttc	cgcgaggtgc	ggattgaaaa	tggctctgctg	ctgctgaacg	1800
gcaagccggt	gctgattcga	ggcgtaaac	gtcacgagca	tcactctctg	catggtcagg	1860
tcattggtga	gcagacgatg	gtgcaggata	tcctgctgat	gaagcagaac	aactttaacg	1920
ccgtgcgctg	ttcgcatatt	ccgaaccatc	cgctgtggta	cacgctgtgc	gaccgctacg	1980
gcctgtatgt	ggtggatgaa	gccaatattg	aaaccacagg	catggtgcca	atgaatcgtc	2040
tgaccgatga	ttccgcgctg	ctaccggcga	tgagcgaaac	cgtaacgcga	atggtgcagc	2100
gcgatcgtaa	tcaccgcag	gtgatcatct	ggtcgctggg	gaatgaatca	ggccacggcg	2160
ctaatacaga	cgcgctgtat	cgctggatca	aatctgtcga	tccttcccgc	ccggtgcagt	2220
atgaaggcgg	cggagccgac	accacggcca	ccgatattat	ttgcccgatg	tacgcgcgcg	2280
tgatgaaga	ccagcccttc	ccggctgtgc	cgaaatggtc	catcaaaaaa	tggctttcgc	2340
tacctggaga	gacgcgccc	ctgatccttt	gcgaatacgc	ccacgcgatg	ggtaacagtc	2400
ttggcggttt	cgctaaatac	tggcaggcgt	ttcgtcagta	tcctcggtta	cagggcgggt	2460
tcgtctggga	ctgggtggat	cagtcgctga	ttaaatatga	tgaaaacggc	aaccggtggt	2520
cggcttacgg	cgggtgatttt	ggcgatacgc	cgaacgatcg	ccagttctgt	atgaacggtc	2580
tggtctttgc	cgaccgcacg	ccgcatccag	cgctgacgga	agcaaaacac	cagcagcagt	2640
ttttccagtt	ccgtttatcc	gggcaaacca	tcgaagtga	cagcgaatac	ctgttccgtc	2700
atagcgataa	cgagctcctg	cactggatgg	tggcgctgga	tggtaaagccg	ctggcaagcg	2760
gtgaagtgcc	tctggatgtc	gctccacaag	gtaaacagtt	gattgaactg	cctgaactac	2820
cgcagccgga	gagcgccggg	caactctggc	tcacagtacg	cgtagtgcaa	ccgaacgcga	2880
ccgatggtc	agaagccggg	cacatcagcg	cctggcagca	gtggcgctctg	gcggaaaacc	2940
tcagtgtgac	gctccccgcc	gcgtcccacg	ccatcccgcga	tctgaccacc	agcgaatgg	3000
atttttgcat	cgagctgggt	aataagcgtt	ggcaatttaa	ccgccagtca	ggctttcttt	3060
cacagatgtg	gattggcgat	aaaaaacaac	tgctgacgcc	gctgcgcgat	cagttcaccc	3120
gtgcaccgct	ggataacgac	attggcgtaa	gtgaagcgac	ccgcattgac	cctaacgcct	3180
gggtcgaaacg	ctggaaggcg	gcgggccatt	accaggccga	agcagcgttg	ttgcagtga	3240
cggcagatac	acttgctgat	gcggtgctga	ttacgaccgc	tcacgcgtgg	cagcatcagg	3300
ggaaaacctt	atztatcagc	cggaaaacct	accggattga	tggtagtgg	caaatggcga	3360
ttaccggtga	tgttgaaagt	gcgagcgata	caccgcatcc	ggcgcggtat	ggcctgaact	3420
gccagctggc	gcaggtagca	gagcgggtaa	actggctcgg	attagggccg	caagaaaact	3480
atcccagacc	ccttactgcc	gcctgttttg	accgctggga	tctgccattg	tcagacatgt	3540
ataccccgta	cgtcttccc	agcgaaaacg	gtctgcgctg	cgggacgcgc	gaattgaatt	3600
atggcccaca	ccagtggcgc	ggcgacttcc	agttcaacat	cagccgctac	agtcaacagc	3660
aactgatgga	aaccagccat	cgccatctgc	tgacgcggga	agaaggcaca	tggctgaata	3720
tcgacgggtt	ccatatgggg	attggtggcg	acgactcctg	gagcccgtca	gtatcggcgg	3780
aattccagct	gagcgccggg	cgctaccatt	accagttggt	ctggtgtcaa	aaataataat	3840

aaccgggcag	gccatgtctg	cccgtatttc	gcgtaaggaa	atccattatg	tactatttaa	3900
aaaacacaaa	cttttggatg	ttcggtttat	tctttttctt	ttactttttt	atcatgggag	3960
cctacttccc	gtttttcccg	atttggctac	atgacatcaa	ccatatcagc	aaaagtgata	4020
cgggtattat	ttttgccgct	atttctctgt	tctcgctatt	attccaaccg	ctgtttggtc	4080
tgctttctga	caaactcgga	acttgtttat	tgcagcttat	aatggttaca	aataaaagcaa	4140
tagcatcaca	aatttcacaa	ataaaagcatt	tttttcaactg	cattctagtt	gtggtttgtc	4200
caaactcatc	aatgtatctt	atcatgtctg	gatcctctag	agtcgacctg	caggcatgca	4260
agctggcact	ggccgctcgt					4279

<210> 2  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic oligonucleotide

<400> 2	
gaatacaaaag cttatgcatg	20

<210> 3  
 <211> 13  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic oligonucleotide

<400> 3	
gaatacaaaag ctt	13

<210> 4  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic oligonucleotide

<400> 4	
aaagcttatg catgcggccg	20

<210> 5  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence

<220>  
 <223> Synthetic oligonucleotide

<400> 5	
cggccgcatc tagagggccc	20

<210> 6  
 <211> 25

<212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Synthetic oligonucleotide  
  
 <400> 6  
 gcggccgcat ctagagggcc cggat 25  
  
 <210> 7  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Synthetic oligonucleotide  
  
 <400> 7  
 aatacaaagc ttatgcatgc ggcc 24  
  
 <210> 8  
 <211> 30  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Synthetic oligonucleotide  
  
 <400> 8  
 aatacaaagc ttatgcatgc ggccgcatct 30  
  
 <210> 9  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Synthetic oligonucleotide  
  
 <400> 9  
 catgcataag ctttgtattc 20  
  
 <210> 10  
 <211> 13  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Synthetic oligonucleotide  
  
 <400> 10  
 aagctttgta ttc 13  
  
 <210> 11  
 <211> 20

<212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Synthetic oligonucleotide  
  
 <400> 11  
 cggccgcatg cataagcttt 20  
  
 <210> 12  
 <211> 20  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Synthetic oligonucleotide  
  
 <400> 12  
 gggccctcta gatgcggccg 20  
  
 <210> 13  
 <211> 25  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Synthetic oligonucleotide  
  
 <400> 13  
 atccgggccc tctagatgcg gccgc 25  
  
 <210> 14  
 <211> 24  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Synthetic oligonucleotide  
  
 <400> 14  
 ggccgcatgc ataagctttg tatt 24  
  
 <210> 15  
 <211> 30  
 <212> DNA  
 <213> Artificial Sequence  
  
 <220>  
 <223> Synthetic oligonucleotide  
  
 <400> 15  
 agatgcggcc gcatgcataa gctttgtatt 30  
  
 <210> 16  
 <211> 1798

<212> RNA  
<213> Unknown

<220>

<223> mRNA sequence for Firefly luciferase

<400> 16

gaauacaaaag	cuaaugcaug	cggccgcauc	uagagggccc	ggauccaaau	ggaagacgcc	60
aaaaacauaa	agaaaggccc	ggcgccauuc	uauccucuag	aggauggaac	cgcuggagag	120
caacugcaua	aggcuangaa	gagauacgcc	cugguuccug	gaacaauugc	uuuuacagau	180
gcacauaucg	aggugaacau	cacguacgcg	gaauacuucg	aaauguccgu	ucgguuggca	240
gaagcuanga	aacgauaugg	gcugaauaca	aaucacagaa	ucgucguaug	cagugaaaac	300
ucucuucaau	ucuuuaugcc	gguguugggc	gccguuauuu	aucggaguug	caguugcgcc	360
cgcgaagcac	auuuauaaug	aaçgugaauu	gcucaacagu	augaacaauuu	cgcagccuac	420
cguaguguuu	guuuccaaaa	agggguugca	aaaaauuuug	aacgugcaaa	aaaaauuacc	480
aaauauccag	aaaauuauua	ucauggauuc	uaaaacggau	uaccagggau	uucagucgau	540
guacacguuc	gucacaucuc	aucuaccucc	cgguuuuauu	gaauacgauu	uuguaccaga	600
guccuuugau	cgugacaaaa	caauugcacu	gauaaugaa	uccucuggau	cuacuggguu	660
accuaagggg	guggcccuuc	cgcauagaac	ugccugcguc	agauucucgc	augccagaga	720
uccuauuuuu	ggcaaucaaa	ucauuccgga	uacugcgauu	uuaaguguug	uuccauucca	780
ucacgguuuu	ggaauguuuu	cuacacucgg	auauuugaua	uguggauuuc	gagucgucuu	840
aauguauaga	uuugaagaag	agcuguuuuu	acgaucccuu	caggauuaca	aaauucaaaag	900
ugcguuugcua	guaccaaccc	uauuuucauu	cuucgccaaa	agcacucuga	uugacaaaau	960
cgauuuaucu	aauuuacacg	aaaauugcuuc	ugggggcgca	ccucuucuga	aagaagucgg	1020
ggaagcgguu	gcaaaacgcu	uccaucuucc	agggauacga	caaggauaug	ggcucacuga	1080
gacuacauca	gcuaauucga	uuacacccga	gggggaugau	aaaccgggcg	cggucgguaa	1140
aguuguucca	uuuuuugaag	cgaagguugu	ggauucggau	accgggaaaa	cgcugggcu	1200
uaucagaga	ggcgaauuau	gugucagagg	accuaugauu	auguccgggu	auguaaaca	1260
uccggaagcg	accaacgccu	ugauugacaa	ggauuggaug	cuacauucug	gagacauagc	1320
uuacugggac	gaagacgaac	acuucuucau	aguugaccgc	uugaagucuu	uaauuaaaau	1380
caaaggauau	cagguggccc	ccgcugaauu	ggaaucgaua	uuguuacaac	acccaacau	1440
cuucgacgcg	ggcguggcag	gucuucccga	cgaugacgcc	ggugaacuuc	ccgccgccgu	1500
uguuguuuug	gagcacggaa	agacgaugac	ggaaaaagag	aucguggauu	acgucgccag	1560
ucaaguaaca	accgcgaaaa	aguugcgcg	aggaguugug	uuuguggacg	aaguaccgaa	1620
aggucuuacc	ggaaaacucg	acgcaagaaa	aaucagagag	auccucauaa	aggccaagaa	1680
ggcgcgaaaag	uccaaaugu	aaaauguaac	uguauucagc	gaugacgaaa	uucuuaagcua	1740
uuguaauccu	ccgagggggc	gagcucccaa	aaaaaaaaaa	aaaaaaaaaa	aaaaaaaaaa	1798

<210> 17

<211> 12

<212> DNA

<213> Consensus Kozack sequence

<400> 17

gccgccgcca	tg	12
------------	----	----

<210> 18

<211> 13

<212> DNA

<213> Consensus Kozack sequence

<400> 18

gccgccrcca	ugg	13
------------	-----	----

<210> 19

<211> 8

<212> RNA

<213> Consensus Shine-Dalgarno sequence

<400> 19

uaaggagg

8